Department of Meteorology, University of Reading, UK, October 29, 2018

Ensemble Data Assimilation

Algorithms – Software – Applications

Lars Nerger

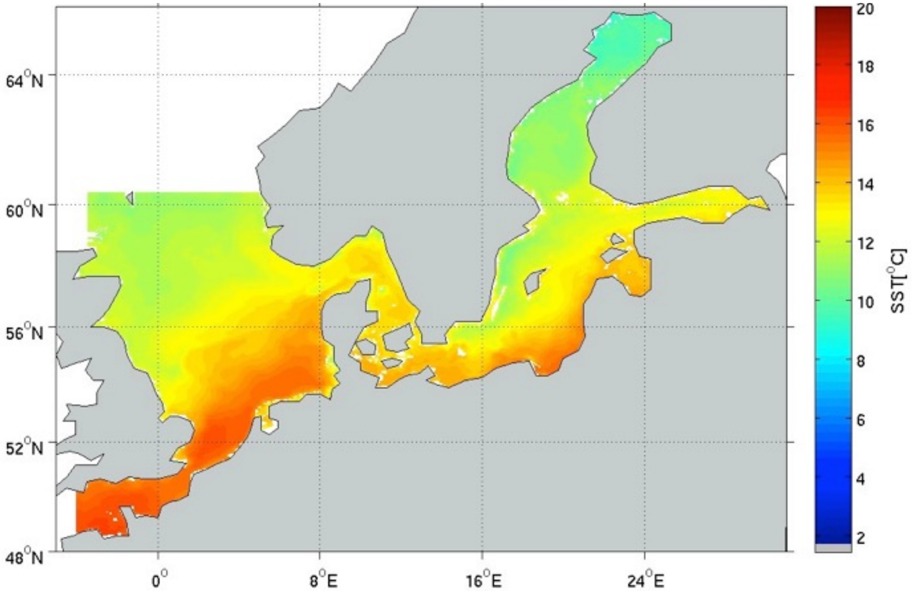
Alfred Wegener Institute for Polar and Marine Research
Bremerhaven, Germany

Acknowledgements:
W. Hiller, J. Schröter, S. Losa, A. Androsov,
H. Pradhan, M. Goodliff, Q. Tang, Q. Yang, L. Mu

ALFRED-WEGENER-INSTITUT
HELMHOLTZ-ZENTRUM FÜR POLAR-
UND MEERESFORSCHUNG

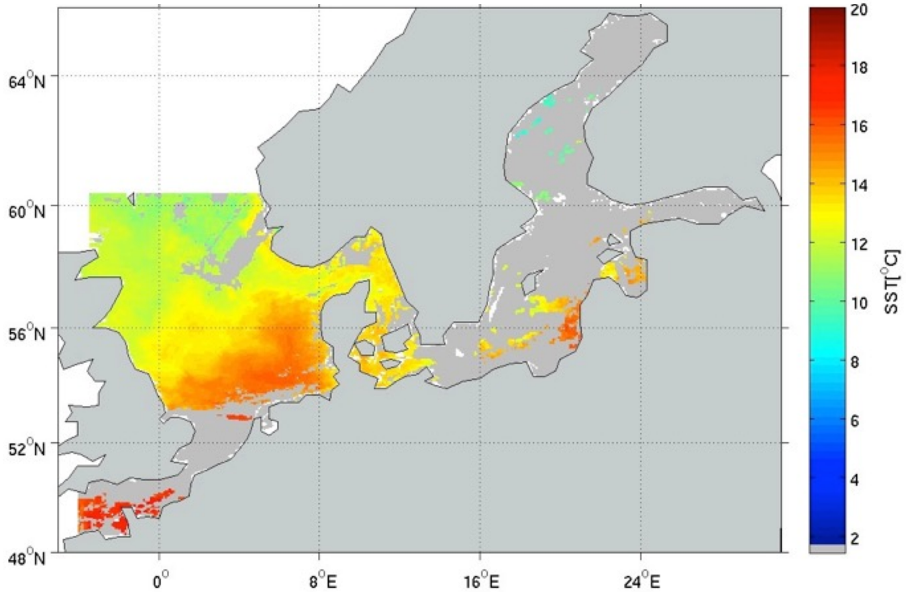
Motivation

Model surface temperature



Information: Model

Satellite surface temperature



Information: Observations

Combine both sources of information

quantitatively by computer algorithm

→ Data Assimilation

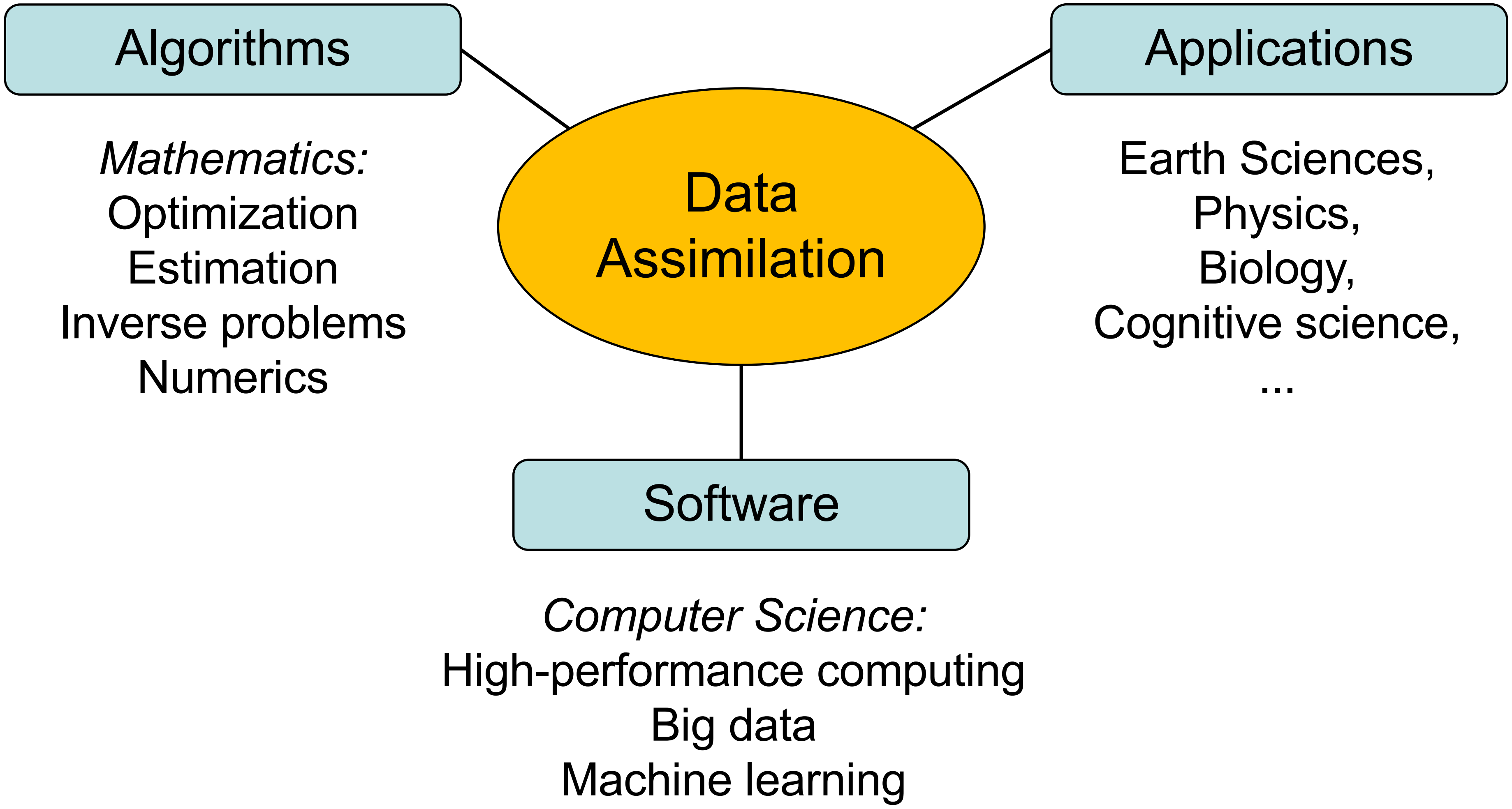
Losa, S.N. et al. *J. Marine Syst.* 105 (2012) 152-162

Data Assimilation

Methodology to combine model with real data

- Optimal estimation of system state:
  - initial conditions (for weather/ocean forecasts, …)
  - state trajectory (temperature, concentrations, …)
  - parameters (ice strength, plankton growth, …)
  - fluxes (heat, primary production, …)
  - boundary conditions and 'forcing' (wind stress, …)
- More advanced: Improvement of model formulation
  - Detect systematic errors (bias)
  - Revise parameterizations based on parameter estimates

Interdisciplinarity of Data Assimilation



Outline

Ensemble Data Assimilation

Algorithms

- Understand behavior of different existing methods
- Develop efficient methods for high-dimensional nonlinear systems

Software

- Make data assimilation easily usable

Applications

- Assess assimilation into realistic model configurations
- Develop methodology for new modeling applications and data types

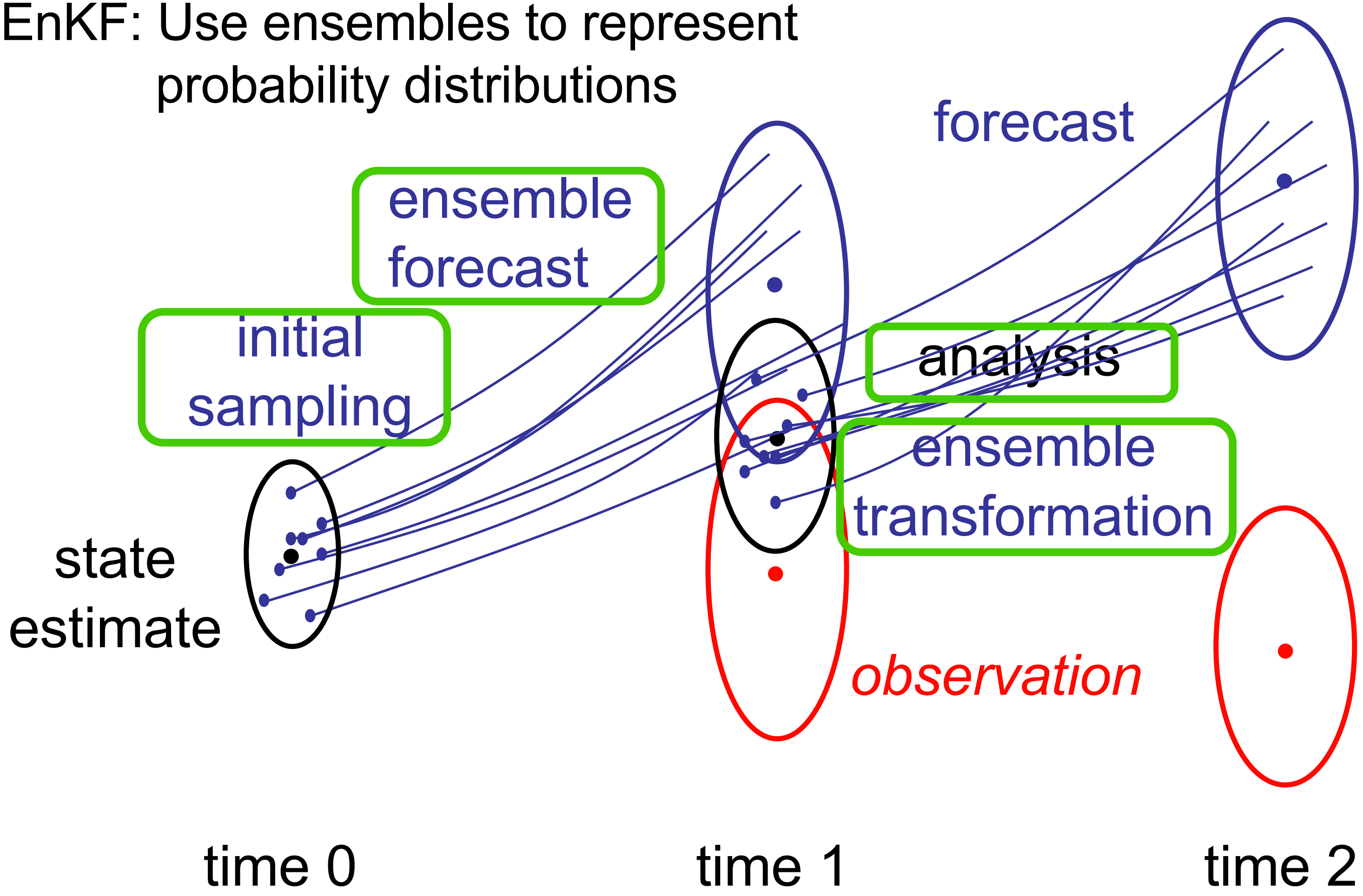
Algorithms

Ensemble-based Filtering

Original EnKF by G. Evensen (J. Geophys. Res. 1994)

Kalman filter: express probability distributions by mean and covariance matrix

EnKF: Use ensembles to represent probability distributions



There are many possible choices!

What is optimal is a research topic

Ensemble-based/error-subspace Kalman filters

A little "zoo" (not complete):

Filter instability
(Nerger 2015)

Which filter should one use?

EnKF(2003)

EnKF(2004)

MLEF

RRSQRT

SPKF

EAKF

ROEK

ESSE

EnKF(94/98)

EnSRF

SEEK

RHF

DEnKF

anamorphosis

Efficiency of SEIK
(Nerger et al. 2005)

SEIK

ETKF

New filter
formulation
(Nerger et al. 2012)

ESTKF

L. Nerger et al., Tellus 57A (2005) 715-735

L. Nerger et al., Monthly Weather Review 140 (2012) 2335-2345

L. Nerger, Monthly Weather Review 143 (2015) 1554-1567

S. Vetra-Carvalho et al., Tellus A 70 (2018) 1445364

Assessing Ensemble Kalman Filters

Mathematical assessment of ensemble Kalman filters limited by

- optimality only proven for Gaussian error distributions
- convergence properties only clear for large ensemble limit

but

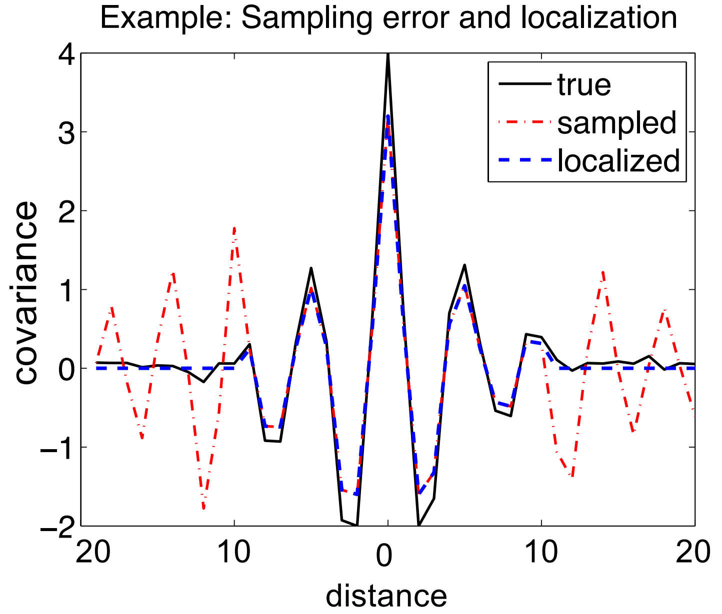
- models are nonlinear -> non-Gaussian distributions
- only small ensemble feasible to run for high-dimensional models

My approach

- compare and characterize behavior of different methods
- reach general conclusions from analyzing differences mathematically

Localization: Why and how?

- Combination of observations and model state based on ensemble estimates of error covariance matrices
- Finite ensemble size leads to significant sampling errors
  - errors in variance estimates
    - usually too small
  - errors in correlation estimates
    - wrong size if correlation exists
    - spurious correlations when true correlation is zero



- Assume: long-distance correlations are small in reality
  - Localization: damp or remove estimated long-range correlations (Houtekamer & Mitchell, 1998, 2001)

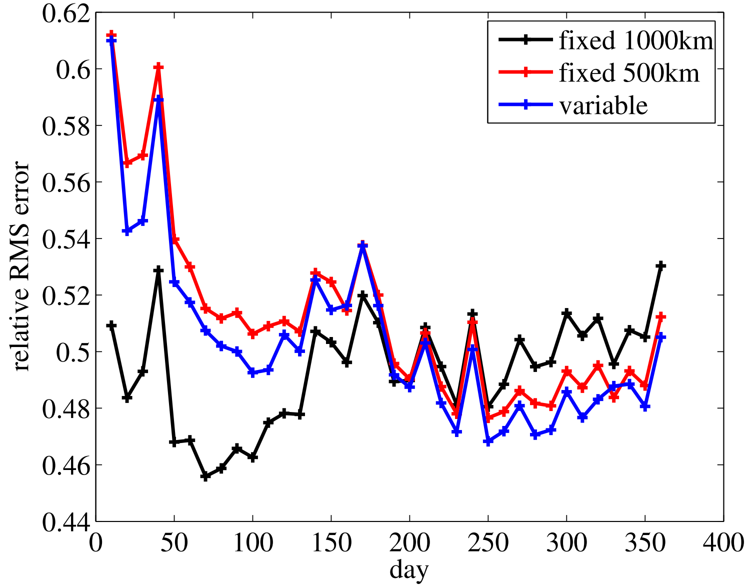
Adaptive localization radius in global ocean model

- Localization radius is usually hand-tuned
- Numerical analysis in small models shows:
  errors minimal when localization radius chosen such that

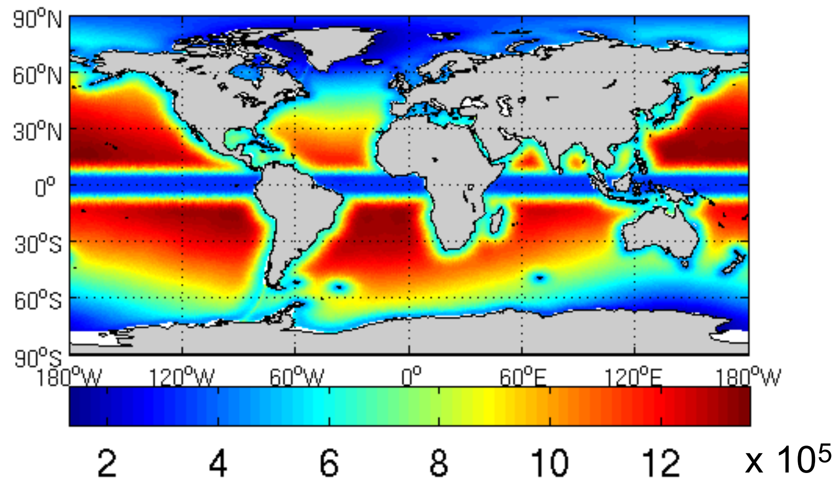
  local sum of observation weights = ensemble size

- Application with FESOM (Finite Element Sea-ice Ocean Model):
  - Fixed 1000km radius leads to increasing errors in 2nd half of year
  - Lower RMS error in sea surface height than fixed 500km radius

Error-reduction of sea surface height



Localization radius [meter]



Kirchgessner, Nerger, Bunse-Gerstner, Mon. Weather Rev., 142 (2012) 2165-2175

Instability of serial observation processing

Two widely used filter categories:

Serial observation processing	Synchronous assimilation
EnSRF, EAKF	***ETKF, SEIK, ESTKF, (EnKF)***
• Perform a loop assimilating each single observation	• Assimilation all observation at a given time at once
• Efficient: Avoids matrix-matrix operations	• Usually using ensemble-space transformations
• Requires diagonal observation error covariance matrix	• Possible for arbitrary observation error covariance matrices
• Localization of state error covariance matrix	• Localization of observation error covariance matrix

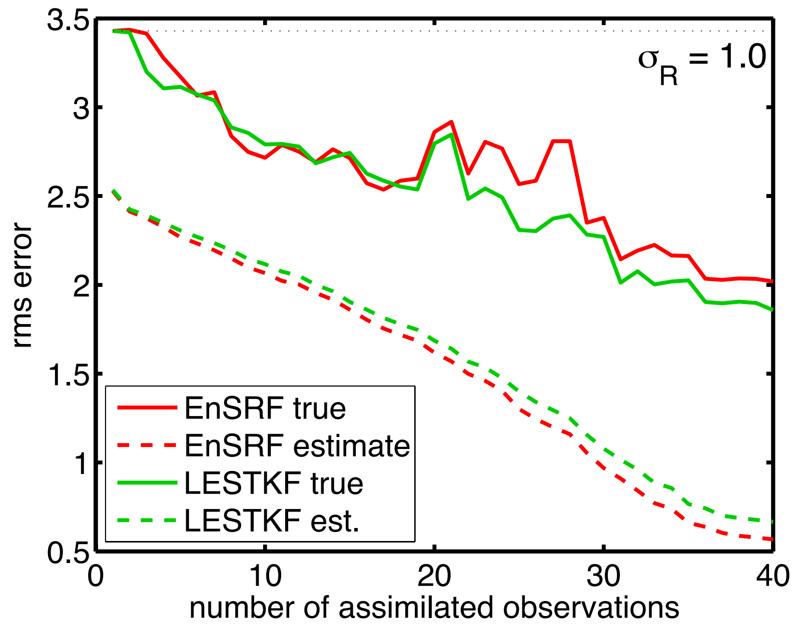
(EnSRF: Whitaker & Hamill, 2002; EAKF: Anderson, 2001)

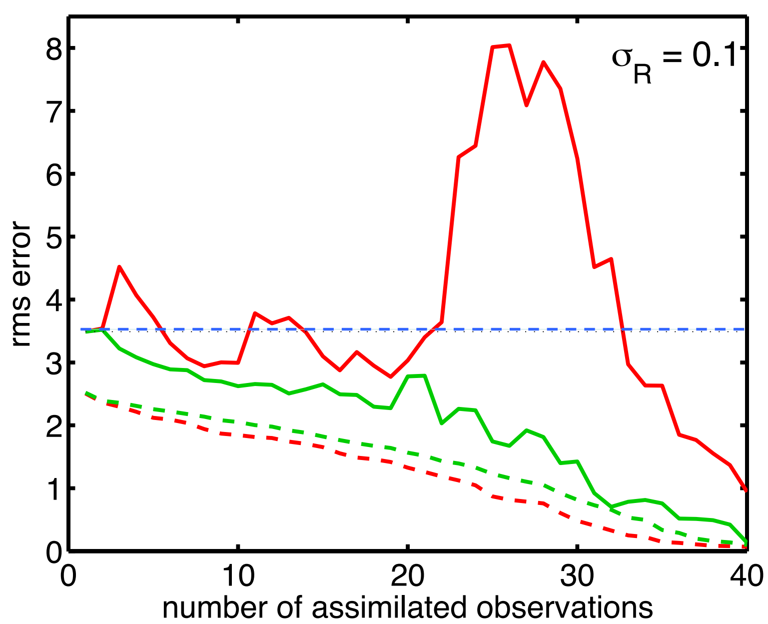
RMS error over number of observations

How does the RMS error develop during the loop over all observations?

Test at first analysis step (Lorenz-96 toy model):

- EnSRF: Compute RMS errors at each iteration
- LESTKF: Do 40 experiments with increasing number of obs.





- Instability leads to larger error for EnSRF in full-length experiments
- Can be relevant in real applications: if observations have locally strong impact

L. Nerger, Monthly Weather Review, 143 (2015) 1554-1567

Inconsistent Matrix Updates

$\mathbf{P}$ State error covariance
$\mathbf{R}$ Obs. error covariance
$\mathbf{H}$ Observation operator

Kalman filter updates covariance matrix according to

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{KH})\,\mathbf{P}^f\,(\mathbf{I} - \mathbf{KH})^T + \mathbf{KRK}^T \quad (1)$$

With Kalman gain

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^T\left(\mathbf{HP}^f\mathbf{H}^T + \mathbf{R}\right)^{-1} \quad (2)$$

this simplifies to

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{KH})\,\mathbf{P}^f \quad (3)$$

Be careful when introducing new adaptions!

(1) and (3) yield same result **only** with gain (2)!

Not fulfilled with localization:

$$\mathbf{K}_{loc} = \left(\mathbf{C}\circ\mathbf{P}^f\right)\mathbf{H}^T\left(\mathbf{H}\left(\mathbf{C}\circ\mathbf{P}^f\right)\mathbf{H}^T + \mathbf{R}\right)^{-1}$$

- Update of **P** is inconsistent in localized EnSRF
  (noted by Whitaker & Hamill (2002), but never further examined)
- Inconsistency also occurs in localized synchronous assimilation ...
  but update is only done once followed by ensemble forecast

Linear and Nonlinear Ensemble Filters

- Represent state and its error by ensemble $\mathbf{X}$ of N states
- Forecast:
  - Integrate ensemble with numerical model
- Analysis:
  - update ensemble mean $\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{X'}^f \tilde{\mathbf{w}}$
  - update ensemble perturbations $\mathbf{X'}^a = \mathbf{X'}^f \mathbf{W}$

  (both can be combined in a single step)

- Ensemble Kalman & nonlinear filters: Different definitions of
  - weight vector $\tilde{\mathbf{w}}$
  - Transform matrix $\mathbf{W}$

ETKF (Bishop et al., 2001)

- Ensemble Transform Kalman filter
  - Assume Gaussian distributions
  - Transform matrix

$$\mathbf{A}^{-1} = (N-1)\mathbf{I} + (\mathbf{HX}'^{f})^{T}\mathbf{R}^{-1}\mathbf{HX}'^{f}$$

  - Mean update weight vector

$$\tilde{\mathbf{w}} = \mathbf{A}(\mathbf{HX}'^{f})^{T}\mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{H}\overline{\mathbf{x}^{f}}\right)$$

(depends linearly on $\mathbf{y}$)

  - Transformation of ensemble perturbations

$$\mathbf{W} = \sqrt{(N-1)}\mathbf{A}^{-1/2}\mathbf{\Lambda}$$

(depends only on $\mathbf{R}$, not $\mathbf{y}$)

NETF (Tödter & Ahrens, 2015)

- Nonlinear Ensemble Transform Filter
  - Mean update from Particle Filter weights: for all particles *i*

    $$\tilde{w}^i \sim \exp\left(-0.5(\mathbf{y} - \mathbf{H}\mathbf{x}_i^f)^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_i^f)\right)$$

    (Nonlinear function of observations $\mathbf{y}$)

  - Ensemble update
    - Transform ensemble to fulfill analysis covariance (like ETKF, but not assuming Gaussianity)
    - Derivation gives

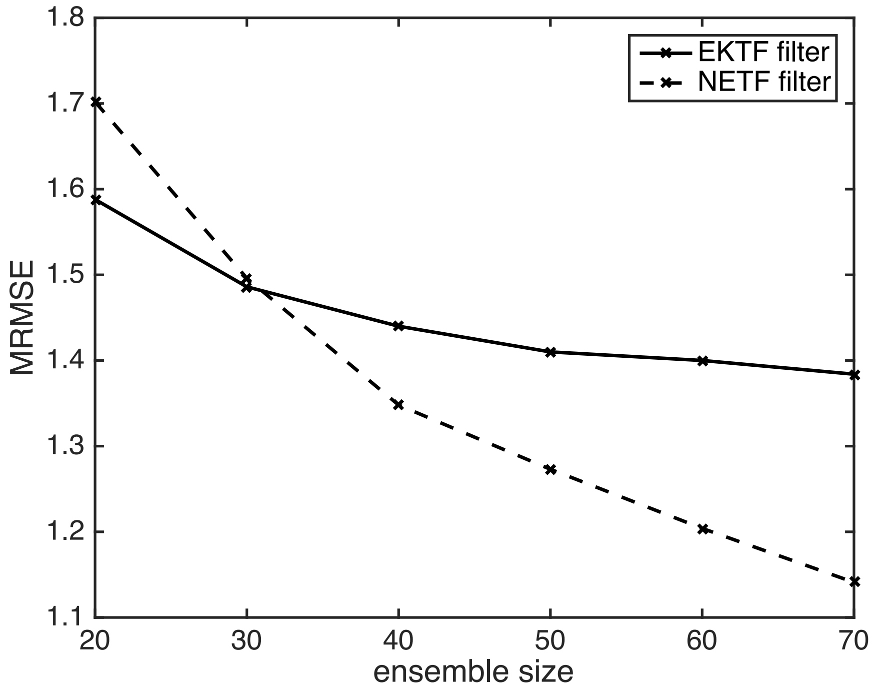
      $$\mathbf{W} = \sqrt{N}\left[\mathrm{diag}(\tilde{\mathbf{w}}) - \tilde{\mathbf{w}}\tilde{\mathbf{w}}^T\right]^{1/2} \Lambda$$

    ($\mathbf{\Lambda}$: mean-preserving random matrix; useful for stability)

Tödter, J. and Ahrens, B. (2015) *Mon. Wea. Rev.* **143**,1347–1367

Performance of NETF – Lorenz-96

- Double-exponential observation errors
- Run all experiments 10x with different initial ensemble



- NETF beats ETKF for ensemble size N > 30
- Larger ensemble needed for Gaussian errors

Kirchgessner, Tödter, Ahrens, Nerger. (2017) *Tellus A*, **69**, 1327766

ETKF-NETF – Hybrid Filter Variants

1-step update (*HSync*)

$$\mathbf{X}^a_{HSync} = \overline{\mathbf{X}}^f + (1-\gamma)\Delta\mathbf{X}_{NETF} + \gamma\Delta\mathbf{X}_{ETKF}$$

- $\Delta\mathbf{X}$: assimilation increment of a filter
- γ: hybrid weight (between 0 and 1; 1 for fully ETKF)

2-step updates

Variant 1 (*HNK*): NETF followed by ETKF

$$\tilde{\mathbf{X}}^a_{HNK} = \mathbf{X}^a_{NETF}[\mathbf{X}^f, (1-\gamma)\mathbf{R}^{-1}]$$

$$\mathbf{X}^a_{HNK} = \mathbf{X}^a_{ETKF}[\tilde{\mathbf{X}}^a_{HNK}, \gamma\mathbf{R}^{-1}]$$

- Both steps computed with increased $\mathbf{R}$ according to γ

Variant 2 (*HKN*): ETKF followed by NETF

Choosing hybrid weight γ

- Hybrid weight shifts filter behavior
- How to choose it?

Possibilities:

- Fixed value
- Adaptive
  - According to which condition?
  - Base on effective sample size $N_{eff} = \sum_i 1/(w^i)^2$

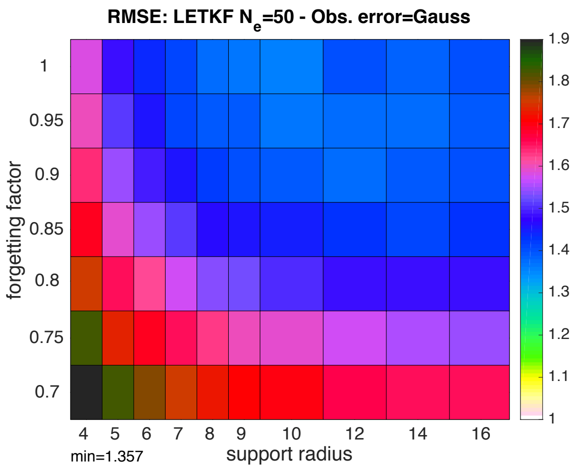
    set

    $$\gamma_{adap} = 1 - N_{eff}/N$$

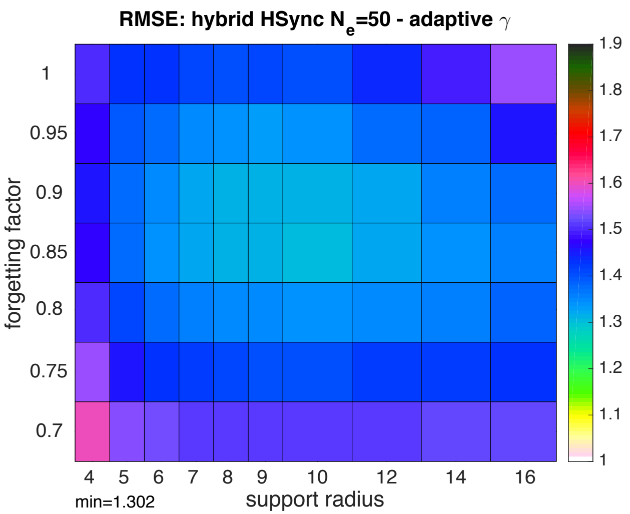
    (close to 1 if N_{eff} small, i.e. small contribution of NETF)

Test with Lorenz-96 Model (ensemble size N=50)

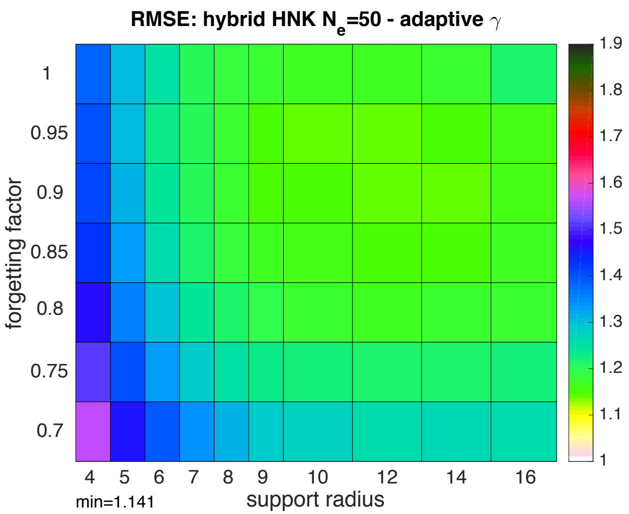
Ensemble size N=50



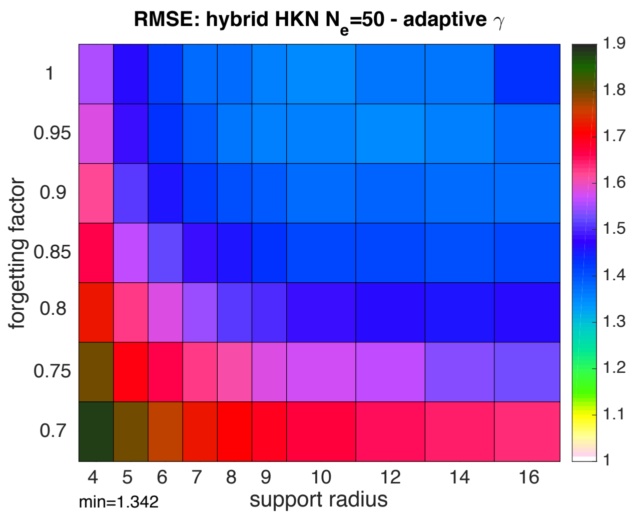
- All hybrid variants improve estimates compared to LETKF & NETF
- Dependence on forgetting factor & localization radius like LETKF
- Similar optimal localization radius
- Largest improvement for variant HNK (NETF before LETKF)
- Currently testing in a larger model ...



4% improvement



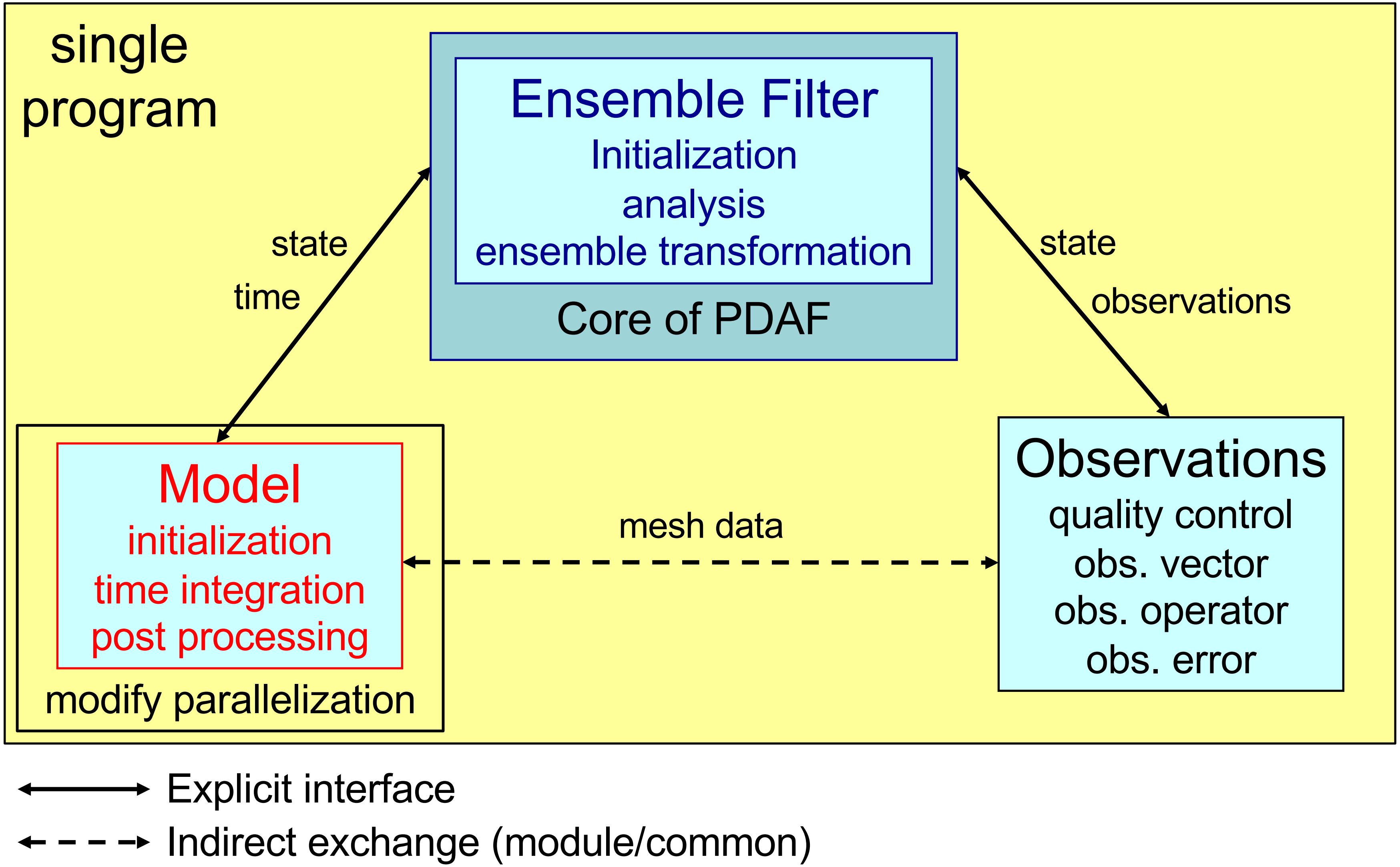
16% improvement



1% improvement

Software

Components of an Assimilation System



L. Nerger, W. Hiller, Computers & Geosciences 55 (2013) 110-118

PDAF: A tool for data assimilation

PDAF — Parallel Data Assimilation Framework

PDAF - Parallel Data Assimilation Framework

- a program library for ensemble data assimilation
- provide support for parallel ensemble forecasts
- provide fully-implemented & parallelized filters and smoothers (EnKF, LETKF, NETF, EWPF … easy to add more)
- easily useable with (probably) any numerical model (applied with NEMO, MITgcm, FESOM, HBM, TerrSysMP, …)
- run from laptops to supercomputers (Fortran, MPI & OpenMP)
- first public release in 2004; continuous further development
- ~310 registered users; community contributions

Open source:
Code, documentation & tutorials at

http://pdaf.awi.de

L. Nerger, W. Hiller, Computers & Geosciences 55 (2013) 110-118

AWI

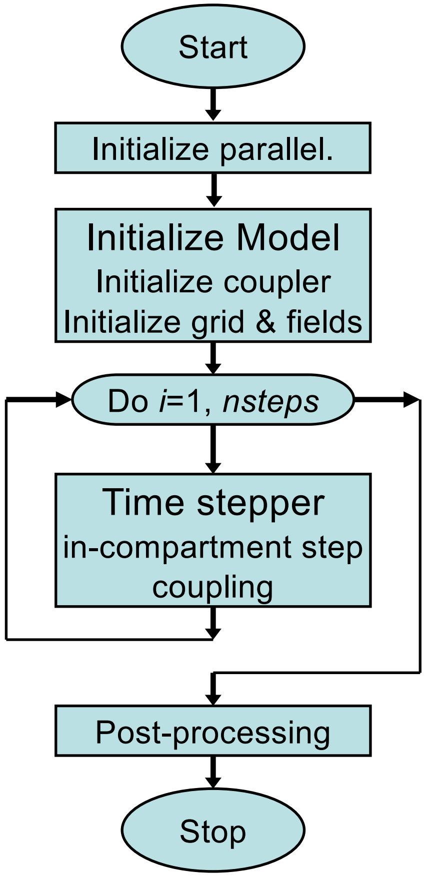
Extending a Model for Data Assimilation

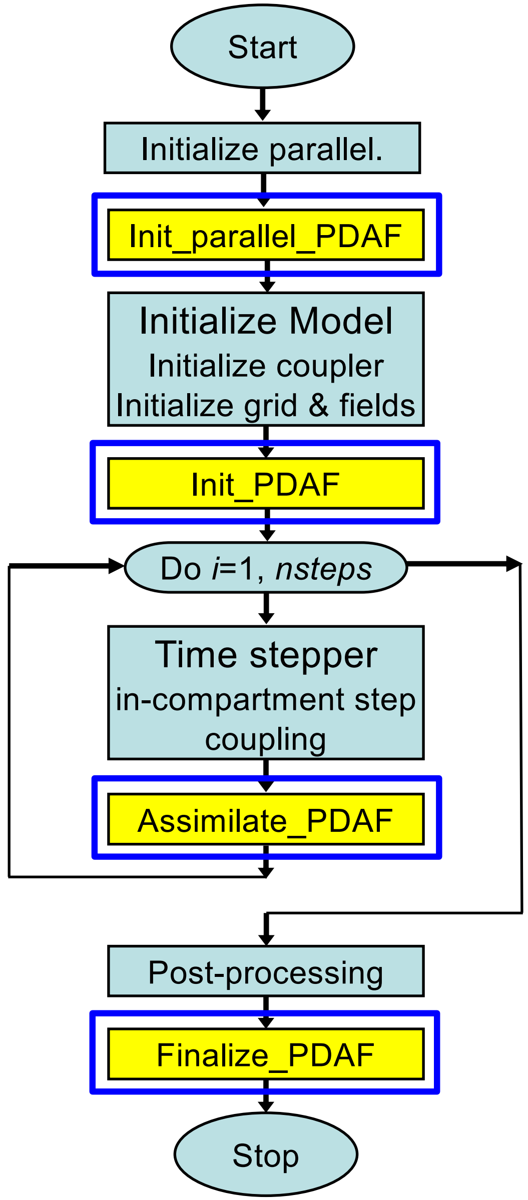
PDAF Parallel Data Assimilation Framework

Model

single or multiple executables

coupler might be separate program





Extension for data assimilation

plus:
Possible model-specific adaption

e.g. in NEMO: treat leap frog time stepping

revised parallelization enables ensemble forecast

(similar to EMPIRE, but more efficient)

PDAF: User-friendliness

Parallel Data Assimilation Framework

Assumption: Users know their model

➜ let users implement assimilation system in model context

For users, model is not just a forward operator

➜ let users extend their model for data assimilation

Keep simple things simple:

- Define subroutine interfaces to separate model and assimilation based on arrays
- No object-oriented programming
(most models don't use it; most model developers don't know it; not many objects would be involved)
- Users directly implement observation-specific routines
(no indirect description of e.g. observation layout)

Example: Value of Efficient Software

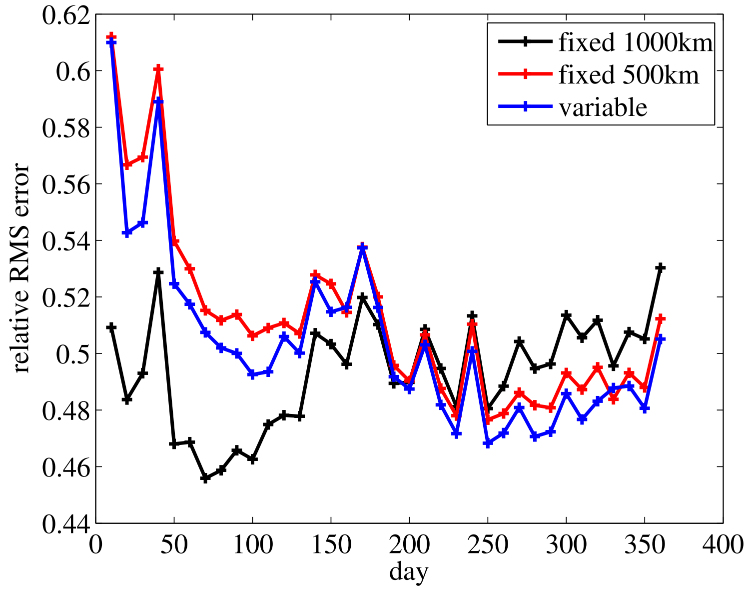
PDAF Parallel Data Assimilation Framework

Adaptive Localization (Kirchgessner et al, 2012)

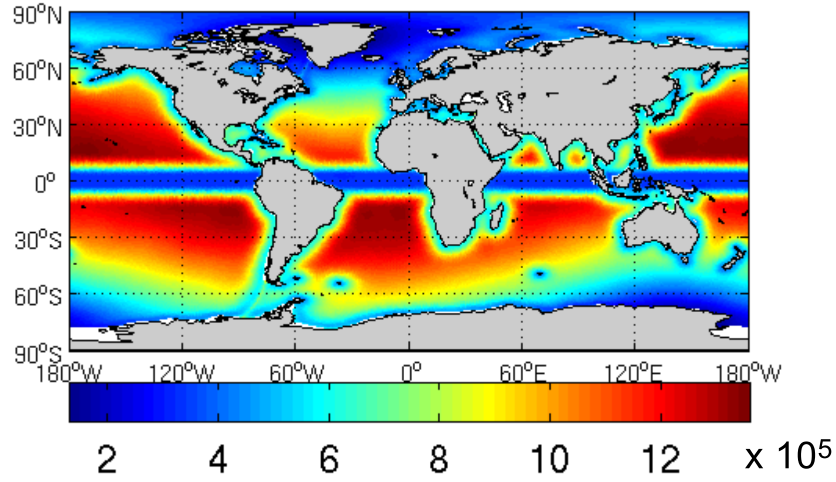
- Original study done with small models (Lorenz-96, shallow water)
- Paper reviewer asked to apply it with full-scale forecast model
- FESOM with PDAF was fully coded without adaptivity
    - ➢ Update PDAF library (just when recompiling)
    - ➢ Adding adaptivity routine and running experiment

} 1 day!

Error-reduction of sea surface height



Localization radius [meter]



Kirchgessner, Nerger, Bunse-Gerstner, Mon. Weather Rev., 142 (2012) 2165-2175

AWI

Applications

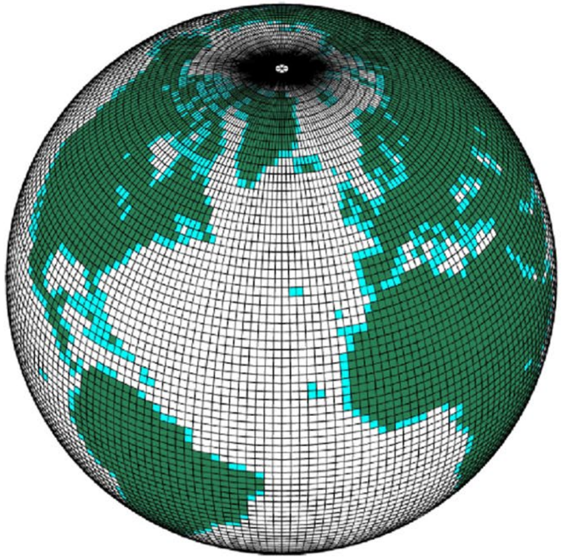
Application Example

Coupled Atmosphere-Ocean Data Assimilation

Qi Tang

Example: ECHAM6-FESOM (AWI-CM)

Atmosphere

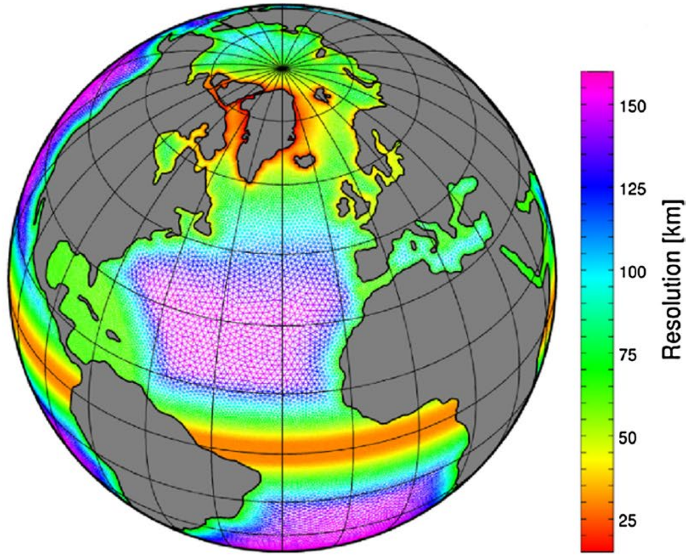


OASIS3-MCT

fluxes

ocean/ice state

Ocean



Atmosphere
- ECHAM6
- JSBACH land

Coupler library
- OASIS3-MCT

Ocean
- FESOM
- includes sea ice

Two separate executables for atmosphere and ocean

Goal: Develop data assimilation methodology for cross-domain assimilation ("strongly-coupled")

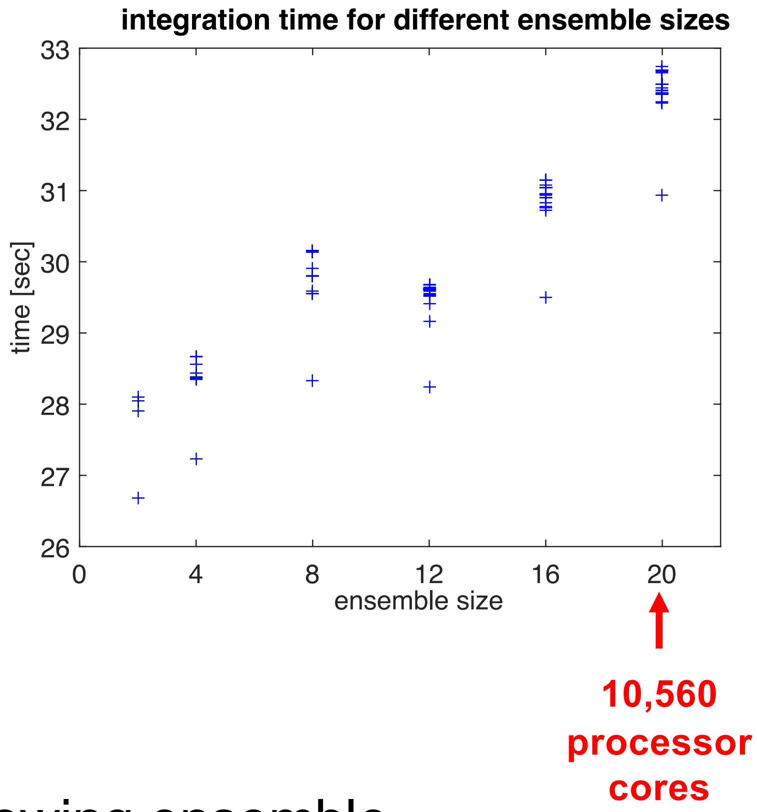
Execution Times (weakly-coupled, DA only into ocean)

MPI-tasks

- ECHAM: 144
- FESOM: 384

Timings (1 day):

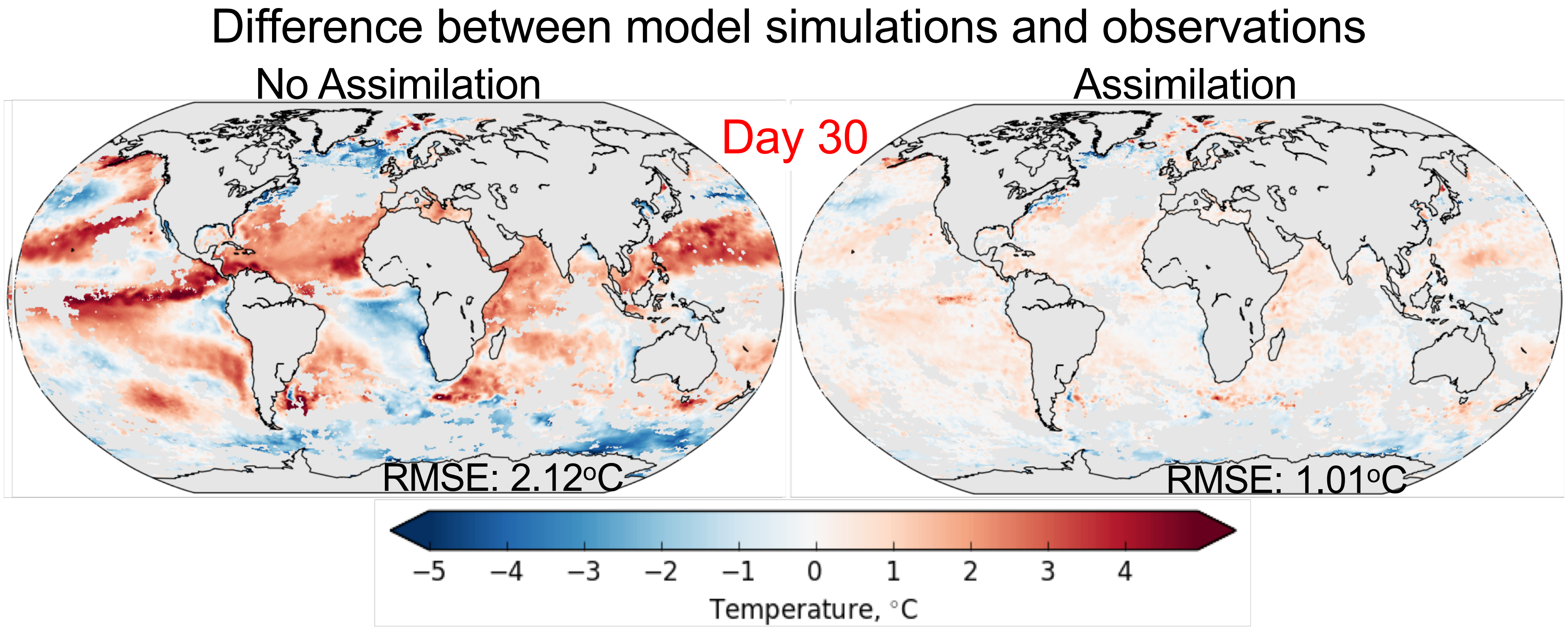
- Ens. forecast: 27 – 33 sec
- Analysis step: 0.5 – 0.9 sec



Scalability:

- Slowly increasing integration time with growing ensemble size (only 16% due to more parallel communication)
- some variability in integration time over ensemble tasks
- Need optimal distribution of programs over compute nodes/racks (here set up as ocean/atmosphere pairs)

Assimilation Effect on Surface Temperature



Surface temperature assimilation successful over 1 year

- Vertical localization required to avoid unrealistic subsurface temperatures

Current work

- Add subsurface profile data (temperature & salinity)
- Assess effect on atmosphere
- Final aim: strongly-coupled assimilation
  (e.g. improve atmospheric state using ocean observations)

Application Example

Assimilation of Satellite Ocean Color Data into Ocean-biogeochemical Model

Himansu Pradhan

IPSO

Coupled <odel: MITgcm - REcoM

MITgcm

General ocean circulation model of MIT (*Marshall et al., 1997*).

Global configuration

80^{o}N - 80^{o}S, 30 layers

Resolution:

lon : 2 deg

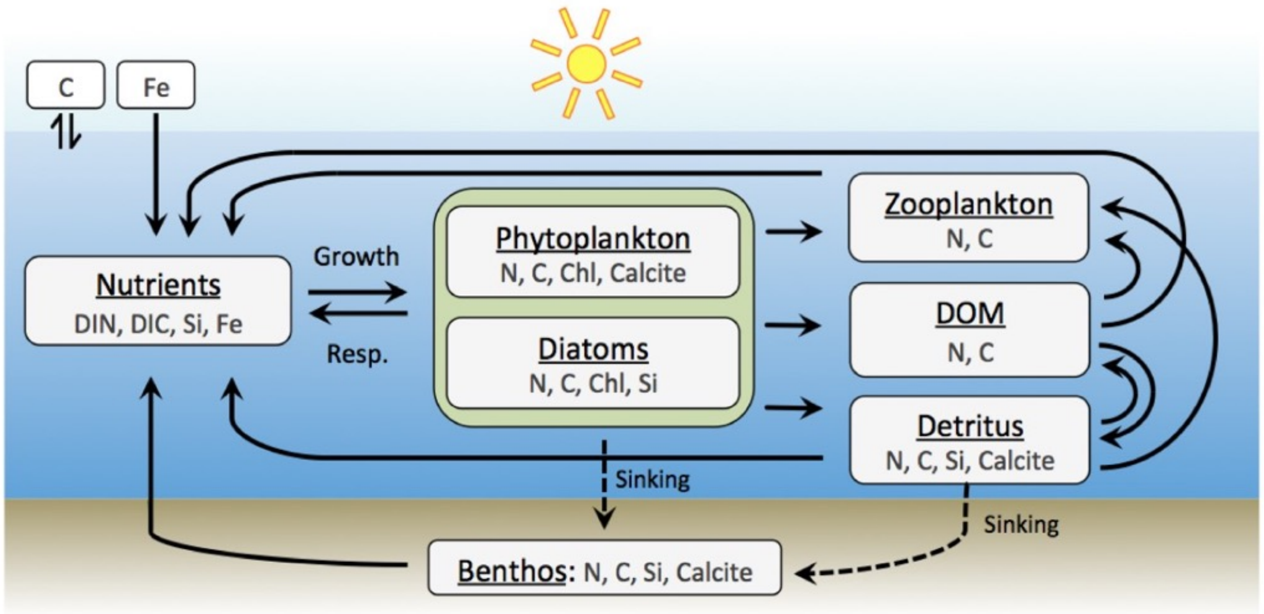
lat : 2 deg in North
up to 0.38 deg in South

layers : 10 m – 500 m

REcoM-2

Regulated **Eco**system **M**odel – Version **2**

(Hauck et al., 2013)

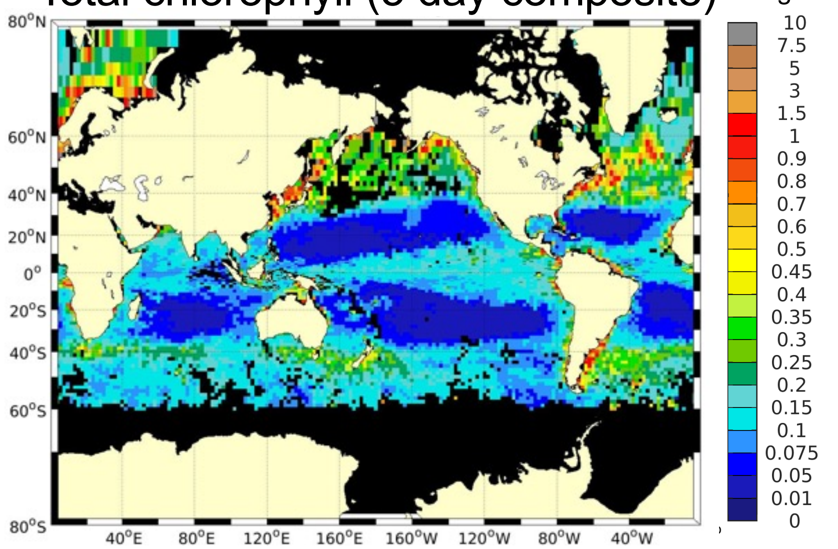


Assimilation of Total Chlorophyll

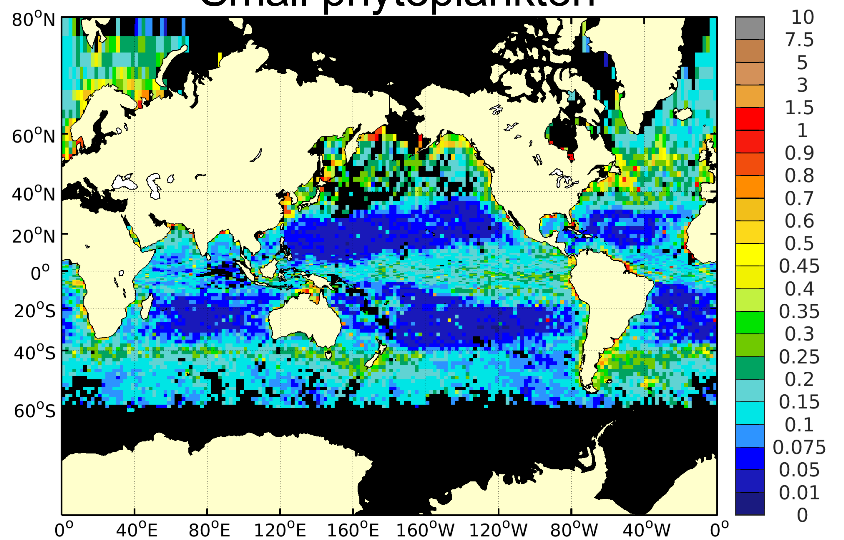
Assimilated:
Total chlorophyll from ESA OC-CCI

Verification: Phytoplankton group data
SynSenPFT (Losa et al. 2018)

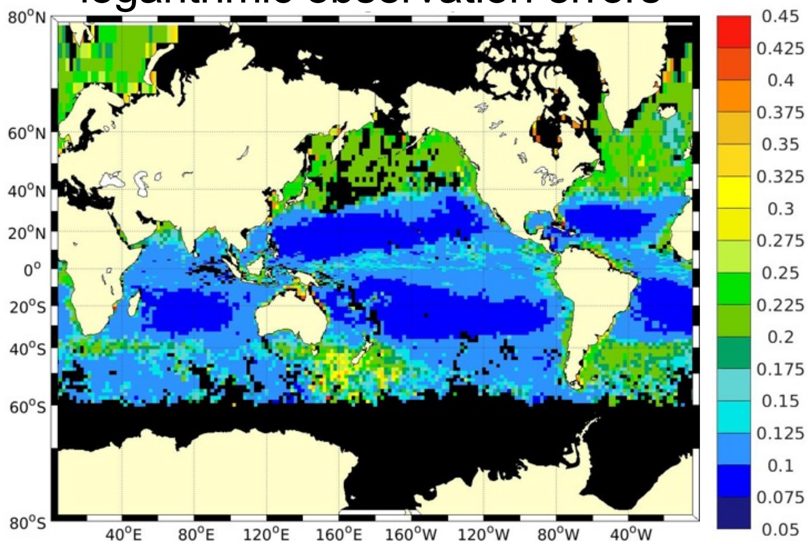
Total chlorophyll (5 day composite)



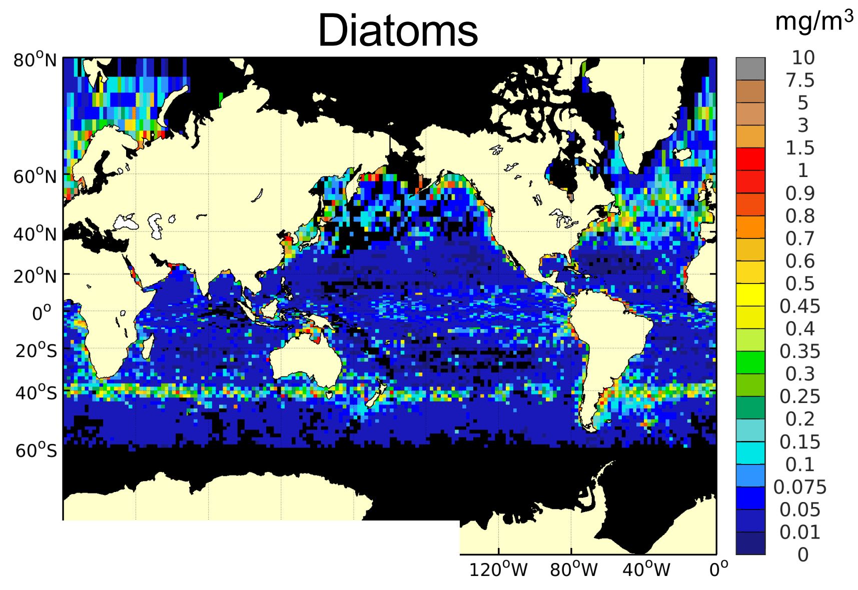
Small phytoplankton



logarithmic observation errors

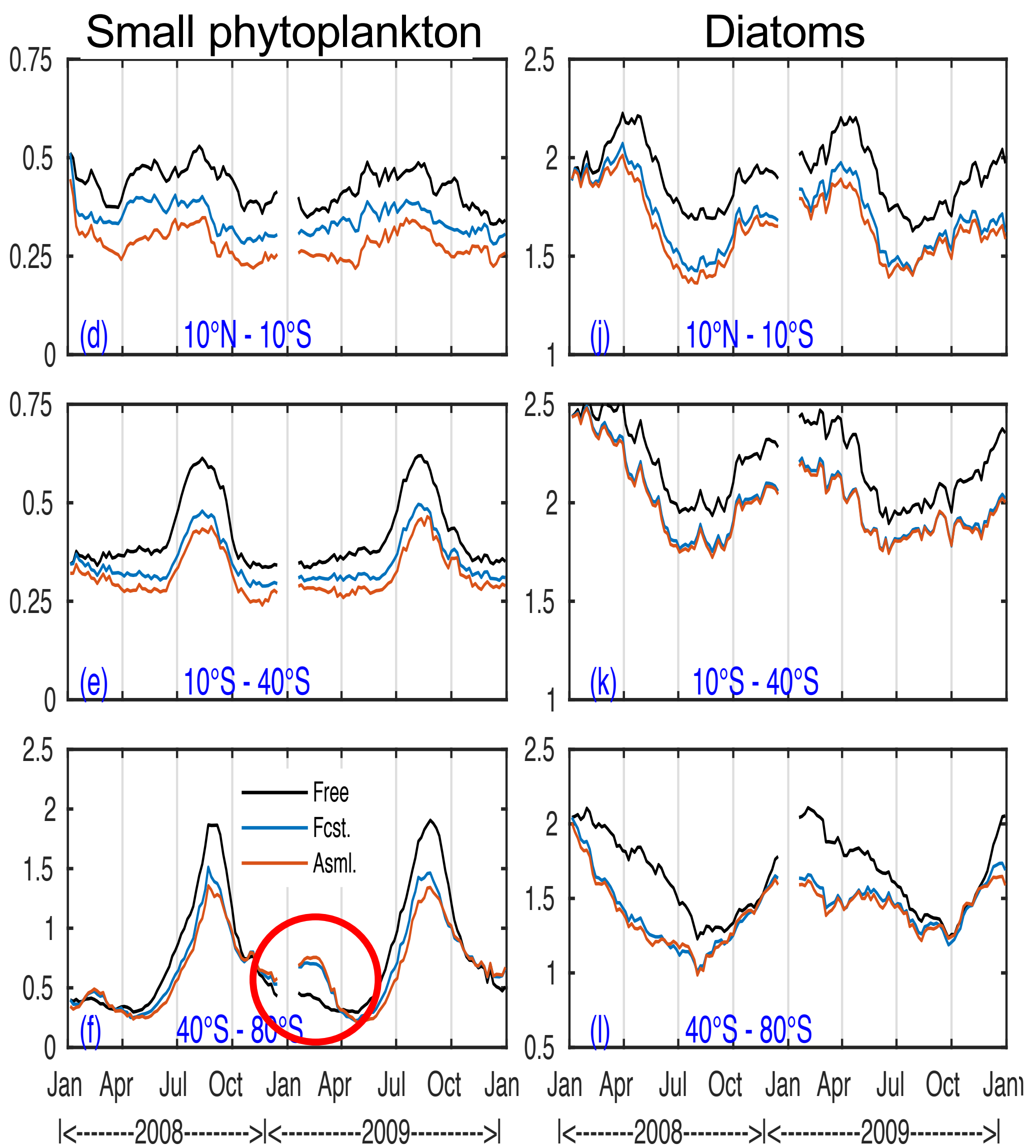


Diatoms



Effect on Chlorophyll in Phytoplankton Groups

logarithmic RMS errors (southern regions)



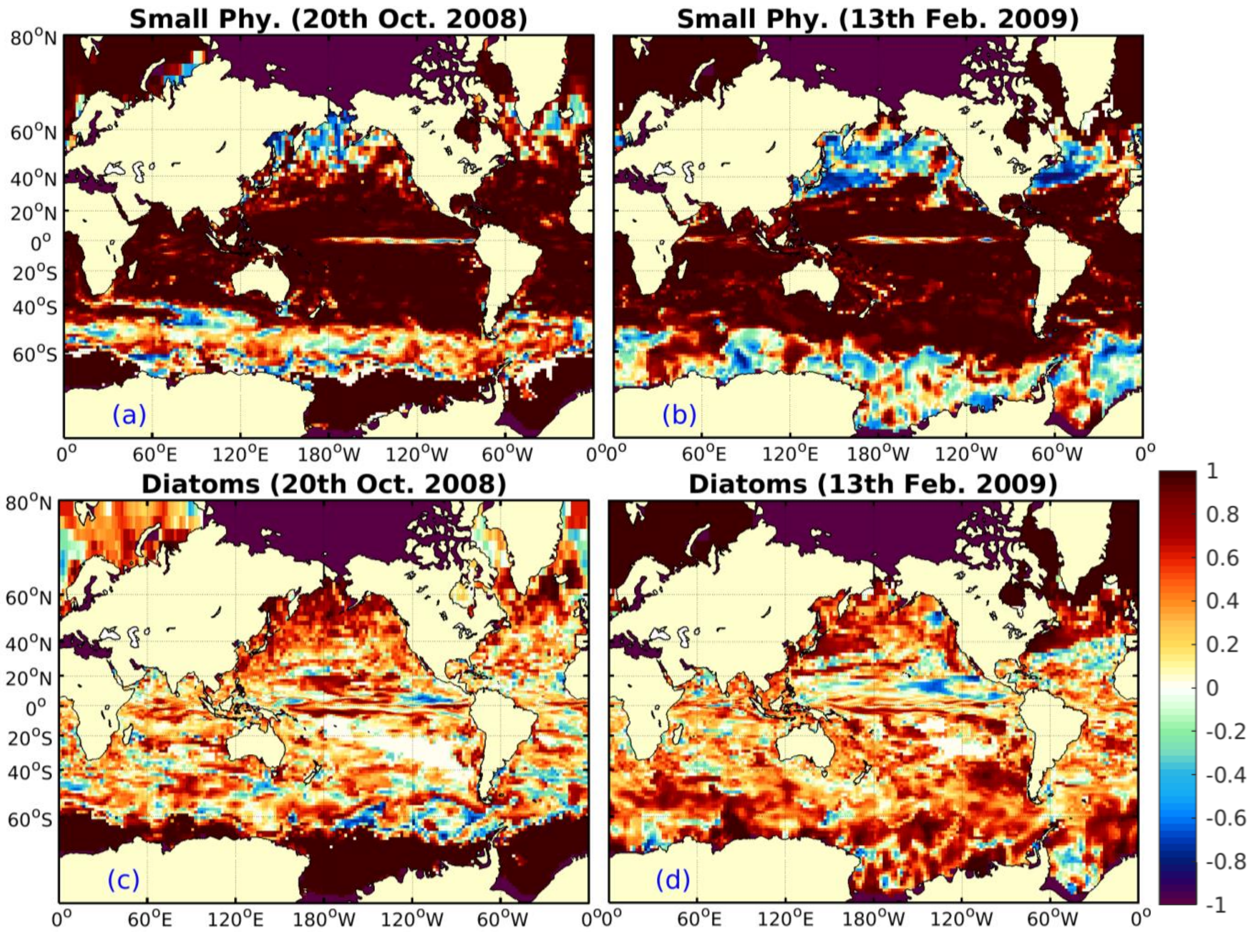
- Assimilation improves groups individually through cross-covariances
- Stronger error-reductions for Diatoms
- Southern Ocean: Particular effect for small phytoplankton at very low concentration

- Current work
  - Asses impact of assimilating chlorophyll group data

Pradhan et al., J. Geophy. Res. Oceans, *under review*

Ensemble-estimated Cross-correlations

Cross correlations between total and group chlorophyll



- Significantly different correlations for small phytoplankton and diatoms
- Negative correlations exist